Normalization Failure in Summation of EXAFS Scans When Gains Change

Matthew A. Marcus

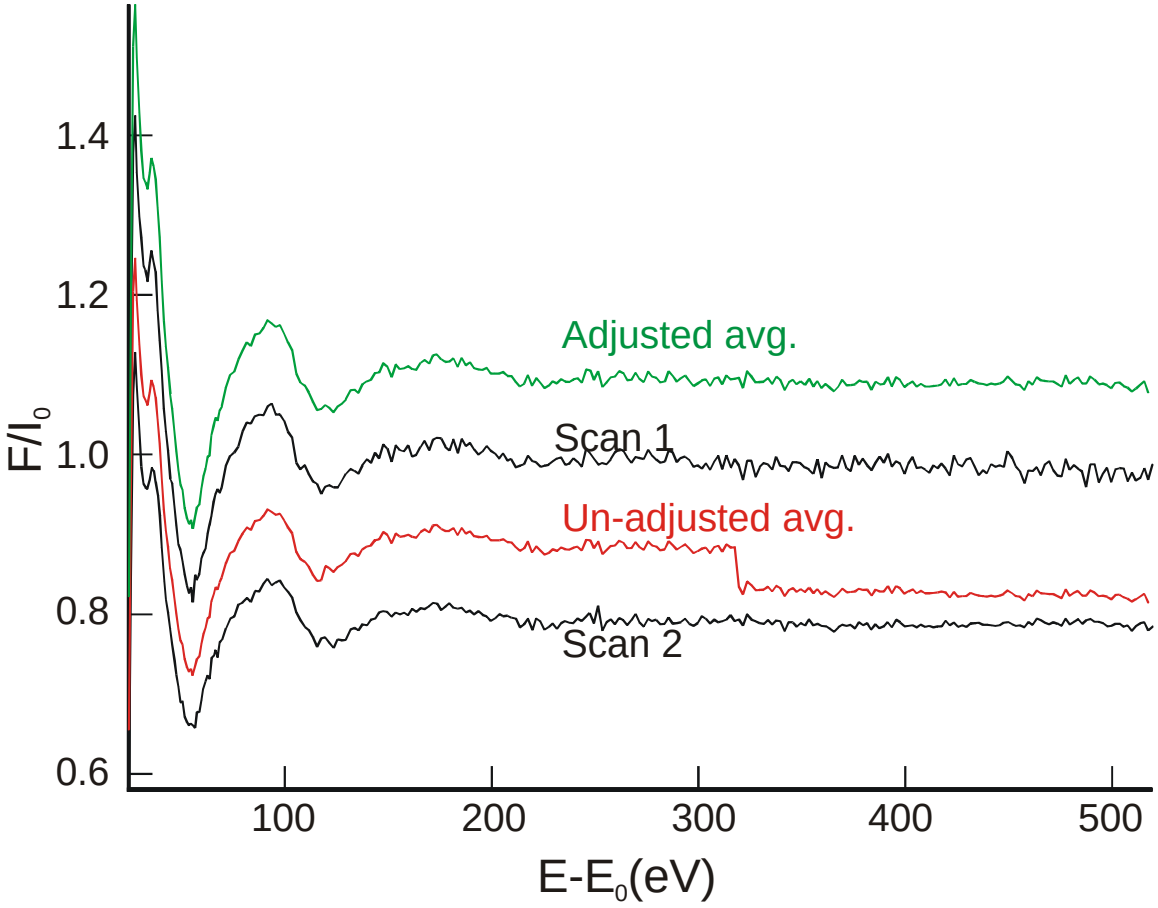
October, 2003

Spectroscopic data collection usually requires dwelling at each energy for many seconds. It is customary to break up that time into many scans, so that drifts, beam dumps, or other problems only affect a small portion of the data. The purpose of this note is to show that under certain conditions, adding up individual scans in the most obvious way can lead to poor data quality.

XAS (X-ray Absorption Spectroscopy) data consist of tables of the readings from several detectors, tabulated against energy. These detectors always include one for the incident beam, I_0 and other channels such as transmission T or fluorescence F. The quantity of interest is a ratio, F / I_0 or a log of a ratio, $\ln \frac{I_0}{T}$. When scans are added up, the usual, and statistically-soundest, procedure is to add up each channel separately, then take the ratio. If the detection sensitivity of the system is constant from scan to scan, so that the ratio for each scan is the same except for noise, then variations in I_0 will normalize out of the average just as for individual scans.

However, a common situation is that the detection system was set up at the end of a fill, and it is found at the start of the next fill that I_0 or a detector channel will saturate. For I_0, the usual remedy is to lower the gain on the monitor so that a given amount of incident beam produces a lower reading. If one adds up the scans taken under both condition, one often finds that fluctuations in I_0 now show up in the ratio, whereas individual scans normalize out perfectly. This note explains why that is and what to do about it.

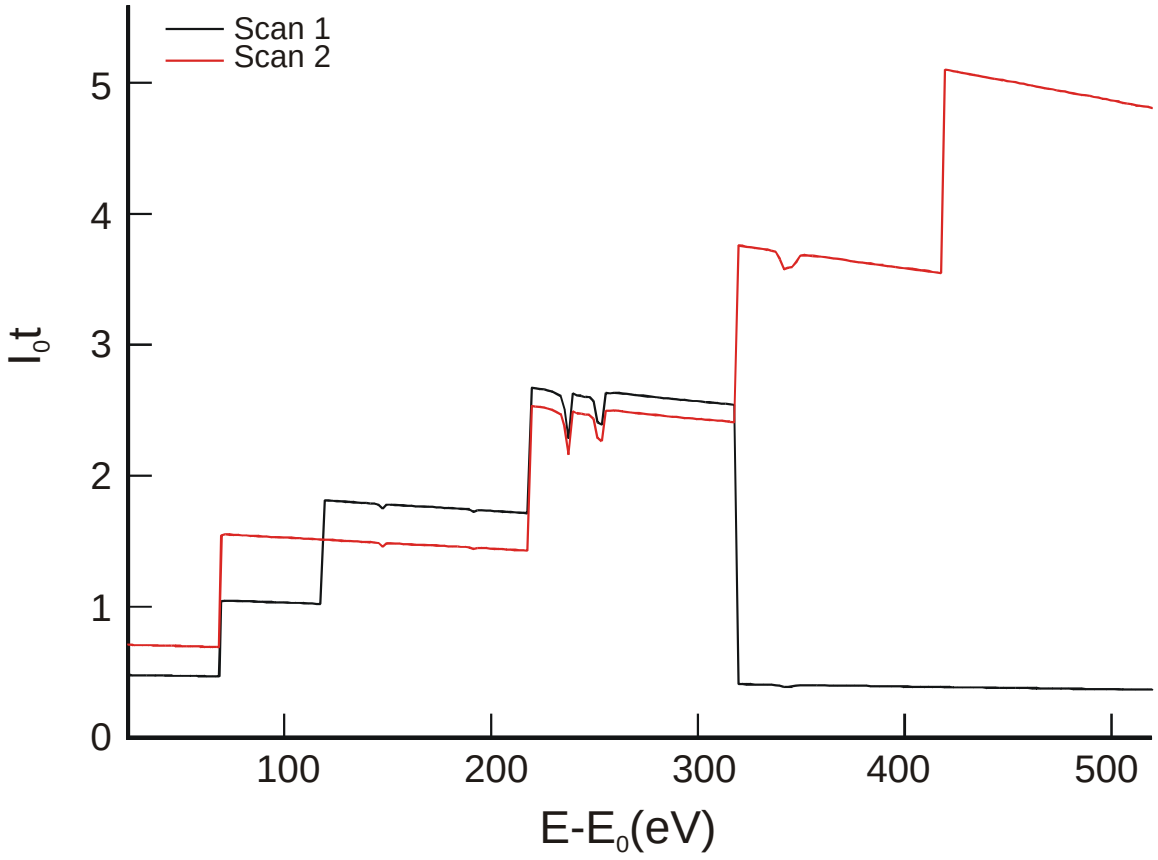
A particularly dramatic example of this problem is shown here:



The two spectra in black are individual fluorescence scans at the Zn edge (thanks to J. Bargar for the use of this data). When the counts in the fluorescence channel of Scan 1 were added to those in Scan 2, then divided by the sum of the I_0 for the two scans, the red curve resulted. When the I_0 for Scan 2 was multiplied by the factor needed to put F / I_0 to the same level as in Scan 1, that is bring the two black curves together, the summed data became as shown in the green curve.

Why the step? The first scan was taken at the end of a fill. In order to make it complete before the beam dump, the count time per point was shortened past where the step occurs in the red curve. The second scan was taken after the new fill. The detector

was moved a little away from the sample, thus resulting in a lower F / I_0 ratio than in Scan 1. The product of count time and I_0 for the two scans is shown here:



As is customary in EXAFS data collection, the time per point increases with distance from the absorption edge, except for the aforementioned drop in Scan 1. While the non-normalizing average effect was stimulated here by a change in count time, it can occur with any difference of I_0 from scan to scan.

Here is an argument as to why we get this effect. Suppose that there is no noise in the spectrum, and that the fluorescence in each of two scans is related to I_0 by the same ratio:

$$F_1(E) = f(E) I_{01}(E)$$
$$F_2(E) = f(E) I_{02}(E) \qquad (1)$$

where the subscripts 1,2 refer to different scans. Now, when we average the two scans, we get

$$\frac{F}{I_0} = \frac{F_1 + F_2}{I_{01} + I_{02}} = \frac{f(E)I_{01}(E) + f(E)I_{02}(E)}{I_{01}(E) + I_{02}(E)} = f(E) \qquad (2)$$

showing that variations in I_0 normalize out. Now, suppose that the sensitivity of the detection channel changes so that instead of $F_2(E) = f(E)I_{02}(E)$, we have $F_2(E) = gf(E)I_{02}(E)$, where g is some constant factor. Think of g as the change in the gain of the detection channel. Now, we have

$$\frac{F}{I_0} = \frac{f(E)I_{01}(E) + gf(E)I_{02}(E)}{I_{01}(E) + I_{02}(E)} = f(E)(1+g/2)\left(1 + (g-1)\frac{(I_{02} - I_{01})}{I_{01} + I_{02}}\right) \qquad (3)$$

showing that the I_0 variations don't normalize if $g \neq 1$. Similarly, if the sensitivity in the I_0 channel changes from scan to scan so that g multiplies I_0, we get the same thing except that g is replaced by $1/g$.

All the above may seem arcane, but the situations I've described happen fairly frequently. Changes in the I_0 gain happen when there is a new fill and the monitor saturates or when the user changes the gas in the I_0 chamber. Changes in fluorescence gain can happen if the beam wanders off the particle, or if one is averaging scans from several spots in order to avoid radiation-damaging any one spot.

What should be done about this problem? If the I_0 gain has changed, then it's easy to make the averaging software multiply the I_0 channel in one or more scans by the appropriate factor effectively to change it back. If the fluorescence gain has changed, then one might think that the fluorescence counts should be multiplied. However, doing

so will over-weight some scans with respect to others in terms of Poisson noise. It is easy to show that when the fluorescence channel is noisier than I_0, then the best signal-to-noise will be obtained by adding up the fluorescence counts in all scans, with no weighting factors. Therefore, instead of multiplying the fluorescence counts by gain-weighting factors, one should multiply I_0 instead. Of course this argument assumes that I_0 is quieter than all other channels, and that the gain change in the fluorescence channel is due to a change in the efficiency of fluorescence detection. A counterexample would be if a Lytle detector or PIN diode were used and the gain on its current amplifier were changed between scans. In this case, the gain change should be compensated for by multiplying the fluorescence readings by the appropriate factor.

To summarize, I have shown that if the detection efficiency of the system changes between scans, then the data should be adjusted to compensate before being summed.